Katz Distributions and Minefield Clearance

Alan Washburn

Department of Operations Research
Naval Postgraduate School
Monterey, CA 93943

INTRODUCTION

The principal uncertainties in clearing a minefield are about the number and type of mines that have not yet been removed. In this paper, we take the point of view that the number of mines should be thought of as a random variable M, and our main object is to show the advantages of a particular "Katz" class of probability distributions for M. We assume throughout that the result of clearance effort is to remove every mine with a known "clearance level" p, independently of all other mines. If $p=1$, we have the case of exhaustive clearance. A theory is hardly required when clearance is exhaustive, and the methods outlined below will be of no use, so we assume $p<1$.

Barring the possibility of exhaustive clearance, any cleared minefield will retain some residual risk to transitors, and quantifying that riskiness should be one of the main goals of theory. Depending on the assessment of risk, it may be advisable to either continue clearance or declare that the minefield is sufficiently cleared that the risk to transitors is bearable. In simplest terms this risk is measured by Simple Initial Threat (SIT), the probability that the first minefield transitor will be killed or damaged by a mine. Since SIT depends strongly on the number of mines remaining, which in turn depends on the number M that were there initially, it is hard to imagine how a basically subtractive

1. REPORT DATE **2006**	2. REPORT TYPE	3. DATES COVERED **00-00-2006 to 00-00-2006**
4. TITLE AND SUBTITLE **Katz Distributions and Minefield Clearance**		5a. CONTRACT NUMBER
		5b. GRANT NUMBER
		5c. PROGRAM ELEMENT NUMBER
6. AUTHOR(S)		5d. PROJECT NUMBER
		5e. TASK NUMBER
		5f. WORK UNIT NUMBER
7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES) **Naval Postgraduate School,Operations Research Department,Monterey,CA,93943-5000**		8. PERFORMING ORGANIZATION REPORT NUMBER
9. SPONSORING/MONITORING AGENCY NAME(S) AND ADDRESS(ES)		10. SPONSOR/MONITOR'S ACRONYM(S)
		11. SPONSOR/MONITOR'S REPORT NUMBER(S)
12. DISTRIBUTION/AVAILABILITY STATEMENT **Approved for public release; distribution unlimited**		
13. SUPPLEMENTARY NOTES **Military Operations Research, Vol. 11, No. 3, 2006, pp. 63-74**		
14. ABSTRACT		
15. SUBJECT TERMS		

16. SECURITY CLASSIFICATION OF:			17. LIMITATION OF ABSTRACT **Same as Report (SAR)**	18. NUMBER OF PAGES **26**	19a. NAME OF RESPONSIBLE PERSON
a. REPORT **unclassified**	b. ABSTRACT **unclassified**	c. THIS PAGE **unclassified**			

clearance activity could result in sufficient knowledge about *M* to support a computation of SIT, unless some information about *M* is input initially. This is in fact the case. According to Bayes theorem, the required information must take the form of a probability distribution for *M*. Katz distributions are a two-parameter class of probability distributions that is broad enough to be practically useful, while simultaneously being narrow enough to permit a significant theory.

Katz distributions have potential applications in areas other than minefield clearance. In considering the reliability of software, for example, one might begin by supposing that there are an unknown number of bugs, *M*, some of which are discovered and removed in the process of using and simultaneously improving the software (Jelinski and Moranda, 1972). In other applications *M* might represent ore pockets, oil deposits, schools of fish or unexploded ordnance. Nonetheless, the language of minefield clearance will be used exclusively below.

The next major section describes Katz distributions, utilizing several subsections and referring to appendices for proofs of theorems.

KATZ DISTRIBUTIONS AND THEIR USES IN MODELING MINEFIELDS

Generalities

Suppose that a region contains an unknown number of mines, *M,* that an action is taken to find and remove the mines, and that *Y* mines are in fact removed by the action. The number of mines that remain is $X \equiv M - Y$. It is *X* that determines the threat of the minefield to subsequent transitors. Even though *Y* is known, *X* is not known exactly because *M* was not known in the first place. Still, the nature of the clearance action taken, together with *Y*, may provide useful information about *X* through an application of Bayes theorem. If the initial distribution of *M* is Katz, then the distribution of *X* will be of the same type, a feature that enables a multi-stage approach to minefield clearance

because the Katz output of one stage can be the Katz input to the next. Katz distributions also have some other appealing properties, so there are good reasons to begin a minefield clearance analysis by assuming a Katz distribution for *M*.

Definition and basic properties

Katz (1965) describes a probability distribution x_0, x_1, ... with the property that

$$x_{j+1}/x_j = \frac{\alpha + \beta j}{1 + j}; \quad j \geq 0. \tag{1}$$

The distribution (1) will be referred to as a "Katz distribution with parameters α and β", provided α and β meet certain restrictions. We will use the notation $M \sim K(\alpha, \beta)$ to express this compactly, with the $\sim$ symbol standing for "is distributed as". Given x_0, equation (1) sequentially determines x_1, x_2, ... Since the sum $x_0 + x_1 + ...$ must be 1, x_0 is determined implicitly.

The parameter α must be nonnegative, since it is the ratio x_1/x_0, and β must be less than 1 to enforce convergence to 0 for large *j*. If $\beta < 0$, then (1) will eventually produce negative probabilities unless $-\alpha/\beta$ is an integer. To prevent this possibility, $-\alpha/\beta$ is required to be an integer when β is negative. The restrictions on parameters are thus that

$$\alpha \geq 0, \quad \beta < 1, \quad \text{and} \quad -\alpha/\beta \text{ is an integer when } \beta < 0. \tag{2}$$

Let the generating function be $g(z;\alpha,\beta) \equiv \sum_{j=0}^{\infty} x_j z^j$. Katz (1965) showed that

$$g(z;\alpha,\beta) = \left[(1-\beta)/(1-\beta z) \right]^{\alpha/\beta}, \tag{3}$$

with (3) being interpreted as exp($\alpha(z$-1)) (the limit as β approaches 0) if $\beta = 0$. It follows that the initial probability must be

$$x_0 = g(0;\alpha,\beta) = (1-\beta)^{\alpha/\beta}, \tag{4}$$

or x_0=exp(-α) if β=0. If M~K(α,β) then (3) implies that

$$E(M) \equiv \mu = \alpha/(1 - \beta) \quad \text{and} \quad Var(M) \equiv \sigma^2 = \alpha/(1 - \beta)^2. \tag{5}$$

It is not hard to establish that a Katz distribution is

- if $\beta < 0$, a binomial distribution with $-\alpha/\beta$ trials and success probability $\beta/(\beta - 1)$,

- if $\beta = 0$, a Poisson distribution with mean α, or

- if $\beta > 0$, a negative binomial distribution. If α/β is an integer, this is the distribution of the number of failures until the α/β th success in a sequence of independent trials where the failure probability is α/β. However, the "number of successes" α/β can actually be any positive real number.

The Katz class includes no other distribution, so it can be thought of as the union of three familiar types.

Since the mean and variance are more familiar parameters than α and β, the solution of (5) for α and β in terms of μ and σ^2 may be useful:

$$\beta = 1 - \mu/\sigma^2 \quad \text{and} \quad \alpha = \mu^2/\sigma^2. \tag{6}$$

Always $\mu \geq 0$ and $\sigma^2 \geq 0$ in (6), but some nonnegative (μ, σ) pairs are impossible because of the restriction that $-\alpha/\beta$ must be an integer when β is negative. This restriction is <u>not</u> imposed by Katz (1965), who simply zeros all probabilities after (and including) the first that (1) would make negative. Unfortunately, this tactic falsifies equations (3) – (6). For example suppose $\alpha = 1$ and $\beta = -2$. Then (1) has $x_1/x_0 = 1$ and $x_2/x_1 = -1/2$, so Katz would take $x_0 = x_1 = 1/2$, $x_i = 0$ for $i \geq 2$. The mean of this distribution is $\mu = 1/2$, not 1/3 as would be obtained by (5). The fact that (3) – (6) are false when $\beta < 0$ and $-\alpha/\beta$ is not an integer is not recognized in Katz (1965), nor in subsequent restatements such as Johnson and Kotz (1969).

Since $\beta = 1 - \mu/\sigma^2$, all (μ, σ^2) pairs where $0 < \mu \leq \sigma^2$ are possible. This covers situations where there is great uncertainty about the number of mines present, as is typically the case in minefield clearance.

The case where $\alpha = \beta = 1$ is a "noninformative prior" in the sense that the ratio x_{j+1}/x_j is 1 for all $j > 0$. It is a limiting case of the negative binomial where all nonnegative

numbers are equally likely and μ approaches infinity. It is not a true distribution because all of the probabilities x_j must approach 0, but may nonetheless serve as a prior distribution for operations such as the one described next.

The Sample-Observe-Subtract (SOS) Property

The main property that makes Katz distributions useful in minefield clearance is that the class is closed under SOS operations. Formally,

<u>Theorem 1</u>:   Let M be the number of mines, suppose $M \sim K(\alpha, \beta)$, let Y be the number of mines removed when each mine is removed with known probability p, independently of the others, and let $X = M - Y$ be the number of mines remaining (not removed). Then, conditional on the event $(Y = y)$ being given, $X \sim K(\alpha', \beta')$, where α' and β' are given by (7) with $q = 1 - p$.

$$\alpha' = q(\alpha + \beta y) \qquad \beta' = q\beta . \qquad\qquad (7)$$

A proof can be found in Appendix A. The same proof can also be found in Washburn (1996), as can proofs of other theorems.

Glazebrook and Boys [1995] introduce a larger class of distributions that is still closed under the SOS operation. Binomial distributions are generalized to "light tailed" distributions, negative binomial distributions are generalized to "heavy tailed" distributions, and the Poisson distribution continues to play its central role. The Katz class can be regarded as a two-parameter subset with additional, convenient analytic properties.

Theorem 1 resolves a certain minefield paradox. Suppose that a minefield is cleared to the .5 level, and that Y mines are removed in the process. One might argue that Y mines must remain, since only half have been removed. But how can it be that the number estimated to remain should increase with the number cleared, since clearance is by its

5

nature subtractive? The paradox disappears when one realizes that clearance to a known level provides both evidence and removal. When $\beta > 0$, the evidence part dominates and the estimated number remaining does indeed increase with the number removed. When $\beta < 0$, the removal part dominates. In the Poisson case $\beta = 0$, the number removed does not affect the distribution of the number that remain.

Since clearance is a process carried out in time, it is likely that clearance times $T_1, \ldots, T_y$ will also be known when $(Y = y)$ is observed. If the magnitudes of these times influence the posterior distribution of M, then the clearance times, as well as the number of mines cleared, should be accounted for. However, there is no effect of this kind as long as the clearance level p is known, regardless of the initial distribution of the number of mines. The proof of this statement can be found in Theorem 2 of Appendix A, and its corollary.

Simple Sampling From a Katz Distribution

Theorem 1 governs the case where Y, the number of mines removed, is observed. There are also circumstances where Y is not observed. One example is where M is the number of mines in region S, but only some fraction q of S (call it S', possibly a transit channel) is of concern. If q is interpreted to be the probability that any given mine will be in S', then the number of mines X in S' is the number remaining after sampling M at the level q, but without observing the results of the sample. Another example is in minefield clearance where the clearance plan is the subject of analyis. Since Y has yet to be observed, any forecast of residual threat cannot be based on Y. Theorem 3 states that X is still Katz, even when Y is not given.

<u>Theorem 3</u>:   Let M be the number of mines, suppose $M \sim K(\alpha, \beta)$, and let X be the number of mines in the sample when each mine is included with probability q,

independently of the others. Then $X \sim K(\alpha', \beta')$, where α' and β' are given by (8) with $p = 1 - q$.

$$\alpha' = \frac{\alpha q}{1 - \beta p} \qquad \beta' = \frac{\beta q}{1 - \beta p}. \qquad (8)$$

A proof can be found in Appendix A. Of course, the number of mines Y *removed* from M is also Katz, but with p and q reversed in (8).

Simple Initial Threat (SIT) for a Katz Distribution

Uncertainty about the number of mines implies uncertainty about whether the minefield is safe for a transitor to cross. The simplest quantification is to define the parameter

$$t \equiv \text{probability that a given uncleared mine kills the transitor}, \qquad (9)$$

and then assume that all mines act independently. For example, suppose that mines are distributed uniformly and independently in a minefield with width W, that each mine actuates with probability B if the transitor's straight line path takes it to within $A/2$ of the mine. Assume also that the transitor will be killed with probability D, conditional on actuation. Then, as long as $W \gg A$ and the transitor's path is near the center of the minefield (ignoring edge effects, in other words), the parameter t is ABD/W. However, t does not need to be calculated in that way – the calculation could involve actuation curves, navigation errors, and edge effects as in Odle (1977).

The transitor is assumed to encounter the mines one at a time. As long as the transitor survives, the probability that the next mine kills it is by assumption t, independently of any others. The probability that all M mines fail to kill the transitor is therefore $(1 - t)^M$, and the probability that the first transitor to enter the minefield is killed is the Simple Initial Threat (SIT):

$$\text{SIT} \equiv 1 - E((1 - t)^M). \qquad (10)$$

If $M \sim K(\alpha, \beta)$, equation (10) can be evaluated by substituting $1 - t$ for z in (3), obtaining

$$\text{SIT} = 1 - g(1 - t;\ \alpha, \beta). \tag{11}$$

If clearance is carried out and Y observed before the transitor enters the minefield, then α' and β' from (7) should be substituted for α and β in (11). If Y has not been observed, then α' and β' from equation (8) should be used instead.

The method of forecasting SIT in current Navy tactical decision aids such as NUCEVL and UCPLN (Wagner, et al, 1999) is based on (7) and (11) with the aforementioned "noninformative prior" for M, essentially a limiting Katz distribution where $\beta=1$ and $\alpha=1$. This distribution is "conservative" in the sense that $E(M)$ is infinite, but such conservatism can have unexpected implications. For example, a minefield cleared to a very low level, with no mines found, would be assessed to have an SIT of nearly 1.

Threat to Transitors after the First

The second and following transitors are much harder to deal with analytically than the first. Odle (1977) gives formulas for several multi-transitor measures, but derivation is non-trivial even when the number of mines is known. An exception is the "catastrophic failure" probability c_n, the probability that none of n transitors is sunk, a concept and term that were introduced by Horrigan (1973). Let Q_n be the catastrophe probability for a single mine. Then c_n is simply Q_n^M for M independent mines. If $M \sim K(\alpha, \beta)$, then the catastrophe probability is

$$c_n \equiv E(Q_n^M) = g(Q_n; \alpha, \beta), \tag{12}$$

where g( ) is again the generating function given by (3). Odle (1977) gives the formula when M is Poisson, a special case. As in the case of SIT, the important thing is that the generating function of M be known.

The single-mine catastrophe probability Q_n would be $(1 - t)^n$ if each transitor's track were chosen independently of the others, but multiple transitors are usually assumed to attempt to follow the same track. In that case the correct "configured" computation of Q_n can become a significant task in itself, particularly if navigation errors are involved, but the degree of difficulty has nothing to do with the distribution of the number of mines present. Regardless of the method used for computing or measuring Q_n, (12) generalizes from one mine to a Katz distributed number of mines.

There appear to be no simple, closed-form formulas other than (12) when multiple transitors are considered, even when the number of mines is known. There are practical methods for calculating the casualty distribution and other statistical measures (Odle, 1977), but the methods do not simplify when the number of mines has a Katz (or even a Poisson) distribution.

A simple upper bound on E_n, the expected number of casualties out of n transitors, can be obtained by observing that the number of casualties cannot exceed M, and therefore that E_n cannot exceed $E(M)$. If each mine causes a casualty with probability at most D whenever it detonates, then a better bound is

$$E_n \leq D \times E(M). \tag{13}$$

If $M \sim K(\alpha, \beta)$, then $E(M)$ is given by (5). Since E_n is necessarily a nondecreasing function of n, (13) is sharpest for large values of n. Of course, $E_1 = \text{SIT}$.

Sums and Partitions of Katz Random Variables

Suppose there are n independent mine populations M_i, with $M_i \sim K(\alpha_i, \beta_i)$, $i = 1, \ldots, n$. Let the total number of mines be $M \equiv M_1 + \ldots + M_n$, and let the clearance level for the ith population be p_i. As usual, take $q_i \equiv 1 - p_i$. Let the number of type i mines cleared be Y_i, with $Y \equiv Y_1 + \ldots + Y_n$, and let the number remaining be X_i, with $X \equiv X_1 + \ldots + X_n$. Of course $X_i + Y_i = M_i$ and $X + Y = M$. These mine populations might be

different kinds of mines in one minefield, the numbers of mines in different minefields, or any other partition of M into n parts. Several questions arise about such mixed minefields

- Does M have a Katz distribution?
- If all of the Y_i are observed, does X have a Katz distribution?
- If only Y is observed, without knowing the mine types, does X_i have a Katz distribution?

Theorem 4 and its corollaries in Appendix A deal with these questions. The answers have a tendency to be a discouraging no, the exceptions being when all of the β_i are equal, or better yet when they are all 0; i.e., when all populations are Poisson.

The answer to the important third question is yes if $Y = 0$, since the observation that $Y = 0$ is equivalent to the observation that $Y_i = 0$ for all i. The answer is also yes if $p = 1$, since in that case $X_i = 0$ for all i. One might hope that the answer would still be yes even if $p < 1$ and $Y > 0$, provided $\beta_i = \beta$ for all i, since the latter condition is sufficient for Y to be Katz (Corollary 1 to Theorem 4). Unfortunately, this is not true. Washburn (1996) shows that conditional independence fails unless $\beta = 0$. This is further evidence that the Poisson case is an especially convenient assumption about the intial number of mines.

If the total number of mines in a minefield is Katz, then can the total be easily partitioned into several independent component Katz distributions? This question might arise because mines can be of different types, or because mines in different parts of a minefield receive unequal clearance effort. The situation is similar to that with summation – the only useful theoretical results are in the case where all of the β_i are equal, especially if they are equal to 0. Theorem 5 of Appendix A summarizes what is known.

AN EXAMPLE

Suppose it is known that approximately 50 mines have been placed in an area with dimensions 5 km long by 2 km wide. Lacking any information to the contrary, the mines are supposed to be all of the same known type, and to be scattered uniformly over the area. It is necessary to clear a channel through the area, but the channel needs to be only 200 m wide, so only about 10% of the mines should be expected in the channel. The top graph in Figure 1 shows the Katz distribution selected for the initial distribution of M, with $\alpha=4.5$ and $\beta=0.1$. It is of the negative binomial type, with a mean of 5 and a standard deviation of 2.48 mines. If (8) is solved for (α,β) with $\alpha'=4.5$, $\beta'=.1$, and $q=.1$, the solution is (23.68, .5263). This Katz distribution for the number of mines in the whole area has a mean of 50 and a standard deviation of 10.27. However, only the mines in the channel are of concern.

We suppose that each mine has a sweepwidth of 20 m against transitors, within which damage is certain. Since the channel is 200 m wide, this corresponds to a threat from each mine of t=.1. The corresponding threat from the mines in the channel, if unswept, is $SIT=g(0.9, 4.5, 0.1)=0.392$ (formula (11)). Suppose it is desired to reduce this threat to .1 by clearing the minefield to some level p. The effect of this sweeping is to change (α,β) to (α', β') according to (8), with $q = 1 - p$, and with (11) subsequently predicting SIT. The required clearance level turns out to be 0.789. This clearance level must be achieved by selecting the number of tracks and runs of each track appropriately, considering the nature of the mines and the sweeping forces. A tactical decision aid such as UCPLN (Wagner, et al, 1999) might be used in planning how to achieve the required clearance level.

Suppose that 10 mines are removed in the process of effecting the clearance plan, a surprisingly large number, given the prior distribution. Given this additional information, the residual threat of the minefield is no longer 0.1. It can be determined by first using

equation (7), with $(\alpha,\beta) = (4.5, .1)$ and $y=10$, to compute $(\alpha', \beta') = (1.1605, .0211)$. The corresponding posterior distribution is shown in the lower part of Figure 1. The posterior SIT from (11) is 0.1112, larger than 0.1 because $\beta>0$ and a large number of mines were found (it is even conceivable that the minefield would be more threatening after clearance than before, although the number of mines found would have to be very large for that to happen). If a threat of 0.1112 is still felt to be too large, then the clearance process must be continued until SIT is sufficiently small.

The clearance process outlined above is sequential in nature, with the need for further clearance depending on the results of clearance to date. This is a realistic feature, since it is characteristic of minefield clearance that the nature of the minefield is determined in the process of clearing it.

This example has been taken from some tutorial notes on mine warfare models that can be downloaded (Washburn, 2005), if desired, along with an accompanying Excel™ workbook (Washburn, 2005). Sheet "Katz" of that workbook incorporates the formulas required to make the above computations.
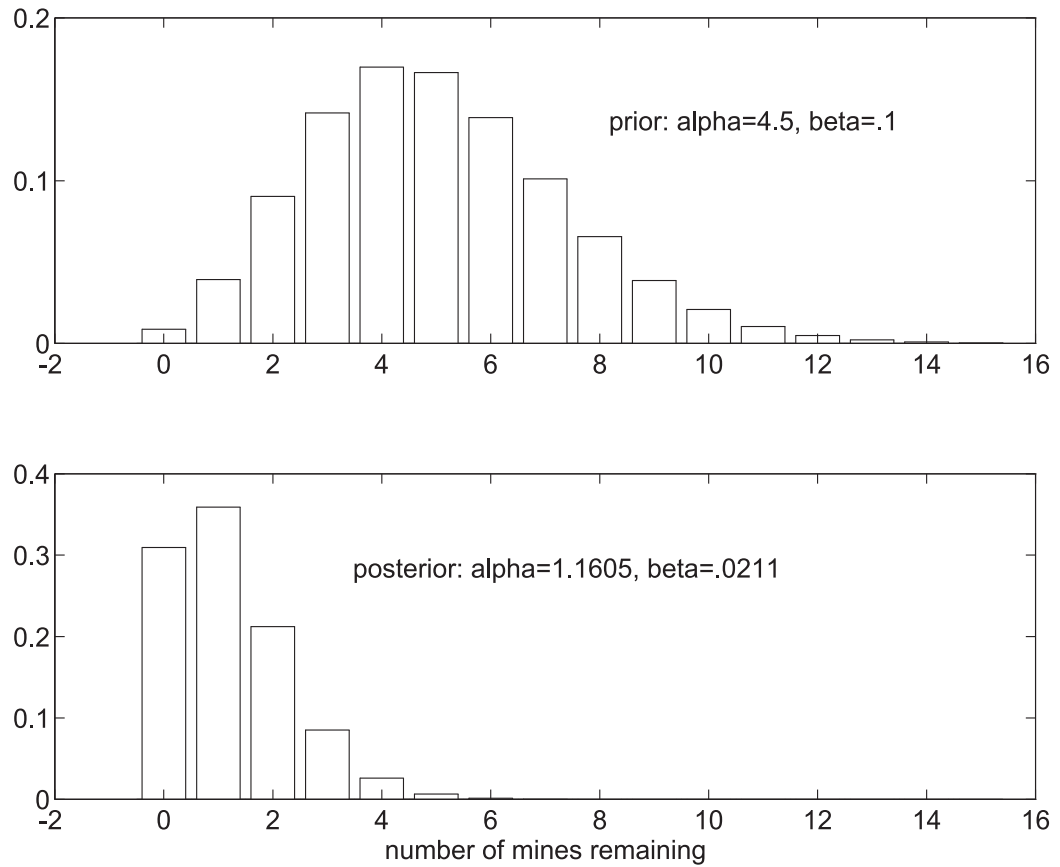
Figure 1: Prior and Posterior (*y*=10) Katz distributions for the example.

BAYESIAN METHODS, KATZ DISTRIBUTIONS, AND DECISION AIDS

As a long-time teacher of Decision Theory at the Naval Postgraduate School, I can assure my readers that it is not easy to convince humans to quantify uncertainty using probability. We are reluctant to do it, even in circumstances of far less importance than clearing a minefield. Bayes Theorem is not intuitive, and posterior distributions are sometimes seen as having more to do with magic than with logic. Given these human tendencies, the nature of current tactical decision aids for minefield clearance is understandable. The user is never asked to quantify expectations about the most important parameter of a minefield — the number of mines that are present. As a result, the effectiveness of clearance must be discussed in terms of clearance level *p*, rather than

the more natural quantity SIT. I have encountered Navy officers who think of the two quantities as opposites; that is, that SIT must be 0.1 if p=0.9. This, too, is natural, albeit wrong, since it is natural to expect decision aids to output quantities that are directly relevant to decision making. But SIT and p are not opposites. Probability theory, correctly applied, is simply unable to determine SIT or any other measure of threat without making some assumption about the number of mines initially present. One can, of course, make a noninformative prior assumption for the initial number of mines, but this is simply getting the user out of the loop by utilizing a (pessimistic) default assumption. One way or another, an assumption is required.

Even if one accepts the idea that the number of mines M must be thought of as a random variable with a prior distribution, it does not necessarily follow that M should be forced to be Katz. The Katz class is closed under some important operations such as SOS, but not under all of them. Particularly when multiple mine types are present, it is possible to make reasonable observations that result in joint posterior distributions not being Katz, or (worse yet) even independent, even when all of the prior distributions are Katz. The Katz class may not be large enough.

Why not permit general distributions for M? A general distribution would require storing 1000 numbers if the maximum conceivable number of mines were 999. A Katz distribution requires only 2, but performing a Bayesian update on a general distribution would nonetheless be trivial with a modern computer. In a different context, NODESTAR (Stone and Corwin, 1995) performs Bayesian updates with 10^6 states, rather than only 10^3. Using a general distribution would also have the advantage that any observation with a known conditional probability law could be the basis of a Bayesian update, which is not true in the Katz case. There seem to be some good arguments for removing all restrictions on the nature of the prior distribution. The idea of using general distributions does not become computationally unwieldy until multiple random variables

must be described jointly. If there were for example five mine types, the number of each of which does not exceed 999, then there would be 10^{15} joint possibilities. Today's computers cannot perform Bayesian updates on that scale, nor will tomorrow's be able to do so. Manipulating general distributions over two or three mine types is currently feasible, and potentially useful, but dealing with 5 mine types is not.

The difficulty with Katz distributions when multiple mines types are present is that there are reasonable observations that destroy the Katz property, in which case one might as well have started with general distributions in the first place. These difficulties largely disappear if all initial mine distributions are independent and Poisson, the special case where $\beta=0$. The decision aid COGNIT (McCurdy, 1987) assumes this. Unfortunately, Poisson distributions have only a single parameter, the mean number of mines, which means that the standard deviation is not independently controllable — a Poisson distribution with mean 100 necessarily has a standard deviation of (only) 10. Inclusion of distributions with $\beta > 0$ (negative binomial distributions) in the permitted class seems essential to model the large uncertainty about mine numbers that is to be expected in practice.

Among the three increasingly general classes of distributions (Poisson, Katz, and general), there are thus serious practical objections that can be made to each. A mine clearance decision aid developer should therefore consider the distribution class question carefully. Here are some further observations in favor of the Katz choice.

Although general distributions for the number of mines are in most cases computationally feasible, it is also true that little is lost by restricting input distributions to be of the Katz type. In fact, the Katz restriction may be operationally welcome, since the entire distribution is determined from only two estimated numbers. With these thoughts in mind, one prototype TDA (MIXER) proposed by the author (Washburn, 1995) employs Katz distributions exclusively, requiring the user to quantify uncertainty

by providing a mean and standard deviation for each mine type. MIXER relies entirely on input files and keyboard responses from the user. Paes (2001) describes a more user-friendly version Javamix with a GUI and some graphics.

The Katz divisibility properties described in Theorem 4 could also prove handy. If a region containing M mines must be divided into two physical parts for two independent clearance operations, then Theorem 4 describes how the numbers of mines in the two parts can be independent, Katz, and still sum to M. The comparable operation in the general case may be difficult or impossible.

Perhaps most important, the availability of an analytic expression for SIT in the Katz case opens up the possibility (as in MIXER) of posing the mathematical problem of minimizing SIT, subject to constraints on the clearance effort, a computational problem that would be much more difficult in the general case. When a variety of clearance resources are available to deal with a variety of mine types, a decision aid capable of producing the "best" clearance plan should be operationally welcome.

In summary,

- Given the central importance of uncertainty about M in minefield clearance analysis, there are some good arguments for taking a Bayesian approach.
- In a Bayesian analysis, Katz distributions are a natural class of probability distributions for the prior distribution of M.

APPENDIX A (Theorems)
The theorems in this appendix are all taken from Washburn (1996), reproduced here for convenience.

Theorem 1 states that Katz distributions are closed under the Sample-Observe-Subtract operation, the fundamental property that makes them useful in mine clearance.

Theorem 1:   Let M be the number of mines, suppose $M \sim K(\alpha, \beta)$, let Y be the number

of mines removed when each mine is removed with probability p, independently of the

others, and let $X = M - Y$ be the number of mines remaining (not removed). Then,

conditional on the event $(Y = y)$ being given, $X \sim K(\alpha', \beta')$, where α' and β' are given

by (10) below with $q = 1 - p$.

Proof: Let $x_j = \Pr(M = j)$ and $x_j^* = \Pr(X = j | Y = y)$; $j = 0, \dots$. Then

$$
\begin{aligned}
x_j^* \Pr(Y = y) &= \Pr(Y = y \cap X = j) \\
&= \Pr(Y = y \cap M = y + j) \\
&= \Pr(Y = y | M = y + j) \Pr(M = y + j).
\end{aligned}
\tag{A1}
$$

But $\Pr(Y = y | M = y + j)$ is the binomial probability of y successes in $y + j$ trials, so,

letting $q = 1 - p$,

$$
x_j^* \Pr(Y = y) = \binom{y + j}{y} p^y q^j x_{y+j}; \quad j = 0, \dots
\tag{A2}
$$

Taking the ratio of successive terms in (A2), the factor $\Pr(Y = y)$ cancels and

$$
x_{j+1}^* / x_j^* = \left\{ \frac{y + j + 1}{j + 1} \right\} q \left\{ \frac{\alpha + \beta(y + j)}{y + j + 1} \right\}.
\tag{A3}
$$

The first { } factor in (A3) is a ratio of combinatorial coefficients, and the second is by

assumption x_{y+j+1}/x_{y+j}. The two $(y + j + 1)$ factors in (A3) cancel, so (A3) is again a linear

function of j divided by $j + 1$, as was to be shown. If α and β satisfy (2), it is easy to

check that the same is true of the revised parameters α' and β', where

$$
\alpha' = q(\alpha + \beta y) \qquad \beta' = q\beta .
\tag{A4}
$$

This concludes the proof of Theorem 1. □

   When the number of mines cleared (Y) is observed, it is likely that the clearance times U_1, …, U_Y will also be observed. These times turn out to have no additional value in making inferences about the initial number of mines M, whether or not M has a Katz distribution, and therefore no value for the residual number of mines $M - Y$. This result may seem counterintuitive. If one searches for 24 hours, finding 5 mines in the first hour and none thereafter, then intuition argues that there are probably no remaining mines, whereas there might be more mines if the clearance times were scattered over the whole clearance period. This intuition might be correct if the probability law $F(\ )$ governing the detection times were unknown, since there is information about $F(\ )$ in the clearance times. If $F(\ )$ is known, however (as it must be if the clearance level is calculable), then the corollary to Theorem 2 states that the clearance times are useless.

Theorem 2:   Let M be a nonnegative random variable, and let T_1, …, T_M be independent, continuous random variables with common distribution function $F(\ )$. Let t be any real number, let I_i indicate the event $(T_i \le t)$, let $Y = I_1 + … + I_M$, and let $\mathbf{U} = (U_1, …, U_Y)$, where U_1, …, U_Y are the nondecreasing order statistics of those T_i for which $T_i \le t$. If m and y are nonnegative integers for which $0 \le y \le m$, and if $\mathbf{u} = (u_1, …, u_y)$ is a real vector, then either $\Pr(Y = y, M = m) = 0$, or

$$\Pr(\mathbf{U} = \mathbf{u}|Y = y, M = m) = \Pr(\mathbf{U} = \mathbf{u}|Y = y). \tag{A5}$$

Proof: Both sides of (A5) are well defined if $\Pr(Y = y, M = m) > 0$. Furthermore, both are 0 unless $u_1 < u_2 < … < u_y < t$, so suppose that those conditions hold. Define the event

$$E_{ym} \equiv (M = m)\bigcap_{i=1}^{y}(T_i = u_i) \bigcap_{i=y+1}^{m} (T_i > t) . \tag{A6}$$

18

Then, since the random variables T_i are all independent by assumption,

$$\Pr(E_{ym}) = \Pr(M = m)\left[\prod_{i=1}^{y} dF(u_i)\right][1 - F(t)]^{m-y}. \tag{A7}$$

The event $(\mathbf{U} = \mathbf{u}) \cap (M = m)$ includes E_{ym} and other mutually exclusive events that have the same probability, since the first y of the T_i are not necessarily the smallest. The number of these events is $y!\binom{m}{y}$, the number of permutations of m things taken y at a time. Thus

$$\Pr(\mathbf{U} = \mathbf{u}, M = m) = y!\binom{m}{y}\Pr(E_{ym}). \tag{A8}$$

Since $\Pr(Y = y, M = m) = \Pr(M = m)\binom{m}{y}F(t)^y[1 - F(t)]^{m-y}$, it is a simple matter to take the ratio $\Pr(\mathbf{U} = \mathbf{u}, M = m)/\Pr(Y = y, M = m)$ to obtain

$$\Pr(\mathbf{U} = \mathbf{u} | Y = y, M = m) = \binom{y}{\mathbf{n}}\prod_{i=1}^{y}[dF(u_i)/F(t)]. \tag{A9}$$

But the right hand side of (A9) does not depend on m, so it must also be $\Pr(\mathbf{U} = \mathbf{u} | Y = y)$.

In other words, conditional on $(Y = y)$ being given, the order statistics $\mathbf{U}$ are distributed as if they were the order statistics of the truncated distribution $F(\ )/F(t)$, sampled y times. $\square$

Corollary: If both sides of (A9) are defined,
$\Pr(M = m | \mathbf{U} = \mathbf{u}, Y = y) = \Pr(M = m | Y = y)$.

Proof: Apply Bayes theorem, recognizing that the right hand side of (A9) does not depend on m. $\square$

According to the corollary, once the number of mines found is known, there is no additional value in observing the times at which they were found.

Theorem 3 states that the residual number of mines is Katz, even if the number of mines removed is not observed.

Theorem 3:   Let M be the number of mines, suppose $M \sim K(\alpha, \beta)$, and let X be the number of mines in the sample when each mine is included with probability q, independently of the others. Then $X \sim K(\alpha', \beta')$, where α' and β' are given by (A13) with $p = 1 - q$.

Proof: Since X is binomial when M is given,

$$E(z^X) = \sum_{j=0}^{\infty} x_j \sum_{i=0}^{j} \binom{j}{i} q^i p^{j-i} z^i \qquad \text{A10)}$$

$$= \sum_{j=0}^{\infty} x_j (qz + p)^j \qquad \text{(A11)}$$

$$= g(qz + p; \alpha, \beta), \qquad \text{(A12)}$$

where $g()$ is the generating function (3). Equation (A11) is obtained from (A10) by combining the factors q^i and z^i, and then employing the Binomial Theorem. Equation (A12) is obtained from (A11) by recalling the definition of the generating function g( ). After rearranging (A12), X can be shown to be Katz with parameters

$$\alpha' = \frac{\alpha q}{1 - \beta p} \qquad\qquad \beta' = \frac{\beta q}{1 - \beta p}. \qquad \text{(A13)}$$

If α and β satisfy (2), then so do α' and β'.   □

Theorem 4 and its corollaries deal with multiple mine types.

Theorem 4:   Assume n independent mine populations M_i, with $M_i \sim K(\alpha_i, \beta_i)$, $i = 1, \ldots, n$. The total number of mines is $M \equiv M_1 + \ldots + M_n$. The clearance level for the ith population is p_i, with $q_i \equiv 1 - p_i$. The number of type i mines cleared is Y_i, with $Y \equiv Y_1 + \ldots + Y_n$, and the number remaining is X_i, with $X \equiv X_1 + \ldots + X_n$. If $\beta_i = \beta$ for all i,

then $M \sim K(\alpha, \beta)$, where $\alpha = \alpha_1 + \ldots + \alpha_n$. Otherwise, M does not have a Katz distribution.

Proof: Since the M_i are all independent, the generating function of M is the product of the individual generating functions (Johnson and Kotz, *op. cit.*, p. 21)

$$g(z) = \prod_{i=1}^{n} \left[(1 - \beta_i)/(1 - \beta_i z) \right]^{\alpha_i / \beta_i} . \tag{A14}$$

If $\beta_i = \beta$ for all i, then (A14) reduces to $g(z) = \left[(1 - \beta)/(1 - \beta z) \right]^{\alpha / \beta}$, the generating function of a Katz random variable. Otherwise, (A14) does not have the required form and M is therefore not Katz. □

Corollary 1:   If $q_i \beta_i = \beta$, $i = 1, \ldots, n$, and if Y_i is observed for $i = 1, \ldots, n$, then

$$X \sim K\left(\sum_{i=1}^{n} (\alpha_i q_i + \beta Y_i), \beta \right).$$

Proof: According to (A4), $X_i \sim K(\alpha_i q_i + \beta Y_i, q_i \beta_i)$ when Y_i is given. Since $q_i \beta_i = \beta$, the conclusion that X has a Katz distribution then follows from Theorem 4. □

Corollary 2:   Suppose $q_i = (1/\beta_i - 1)/(1/\beta - 1)$ for some parameter β; $i = 1, \ldots, n$. Then $X \sim K(\alpha_{\text{TOT}}, \beta)$, where

$$\alpha_{\text{TOT}} = \sum_{i=1}^{n} \beta(\alpha_i / \beta_i) . \tag{A15}.$$

Alternatively, if $p_i = (1/\beta_i - 1)/(1/\beta - 1)$ for $i = 1, \ldots, n$, then $Y \sim K(\alpha_{\text{TOT}}, \beta)$. If $\beta_i = 0$ for all i, then take β/β_i to be q_i (for X) or p_i (for Y) in (A15).

Proof: The condition on q_i enforces $\beta_i' = \beta$ and $\alpha_i' = \beta(\alpha_i/\beta_i)$ in (A14), which applies when the number of mines cleared is not observed. The conclusion then follows from

Theorem 4. If the condition on p_i holds, then the same logic applies to Y, the number of mines not removed. $\square$

Theorem 5 deals with methods for sampling a mixed minefield when the Katz components $M_1,\ldots,M_n$ are all independent. One could, of course, sample n Katz random variables, but it will generally be more efficient to sample M (the total), and then divide it into n parts. Theorem 5 relies on the fact that a certain distribution is of the MCK type, as explained below.

If $M \sim K(\alpha, \beta)$, then a closed-form expression for the probability mass function of M, valid if $\beta < 0$ or $\beta > 0$, is

$$P(M = m) = \frac{(-\alpha/\beta)_m}{m!}(1-\beta)^{\alpha/\beta}(-\beta)^m; \quad m \geq 0. \tag{A16}$$

The notation $(x)_m$ is taken from Feller (1957) where $(x)_m$ is defined to be $x(x-1)\ldots$ $(x - m + 1)$ for $m \geq 1$, with $(x)_0 \equiv 1$ (m is a nonnegative integer, but x can be any real number). The limit as $\beta \to 0$ produces a Poisson distribution, so in that sense (A16) is valid for all (α, β) satisfying (2). If $M_i \sim K(\alpha_i, \beta)$, and if $M_1, \ldots, M_n$ are all independent, then $M \sim K(\alpha, \beta)$ according to Theorem 3. Let $\mathbf{M} \equiv (M_1, \ldots, M_n)$, and $\mathbf{m} \equiv (m_1, \ldots, m_n)$. Then

$$P(\mathbf{M} = \mathbf{m}|M = m) = \frac{\prod_{i=1}^{n} P(M_i = m_i)}{P(M = m)}; \quad m_i \geq 0, m = m_1 + \ldots + m_n. \tag{A17}$$

All of the factors involving $(1 - \beta)$ and $(-\beta)$ raised to powers cancel in (A17), leaving

$$P(\mathbf{M} = \mathbf{m}|M = m) = \frac{\prod_{i=1}^{n} \dfrac{(-\alpha_i/\beta)_{m_i}}{(m_i)!}}{\dfrac{(-\alpha/\beta)_m}{(m)!}}; \quad m_i \geq 0, m = m_1 + \ldots + m_n. \tag{A18}$$

The distribution (A18) will be referred to as a "multivariate conditional Katz distribution with parameters $\boldsymbol{\alpha}$, β, m and n", or MCK for short. The MCK distribution is a multivariate hypergeometric distribution when $\beta < 0$, or a multinomial distribution in the limit as $\beta \to 0$ (Johnson and Kotz, *op. cit.,* p. 281). When $\beta > 0$, the MCK distribution has been called a multivariate Polya-Eggenberger distribution (Johnson and Kotz, 1977) on account of its relationship to certain urn-sampling schemes, or the multivariate Polya distribution (Janardin and Patil, 1970). Thinking of M_i as the number of balls in an urn leads to a practical way of generating $\mathbf{M}$ in a Monte Carlo simulation, since only a single Katz sample of the total M is really required. This is the gist of Theorem 5.

Theorem 5: Let $M \sim K(\alpha, \beta)$, where $\alpha = \alpha_1 + \ldots + \alpha_n$, $\alpha_i \geq 0$ for $1 \leq i \leq n$. The pair (α_i, β) is assumed to satisfy (2) for $1 \leq i \leq n$. Consider the following procedure for placing M balls in n urns. For $k = 0, \ldots, M - 1$, the $k + 1$st ball is placed in urn i with probability p_i, where

$$p_i = \frac{\alpha_i + \beta k_i}{\alpha + \beta k}, \tag{A19}$$

and where k_i is the number of balls already in urn i. If M_i is the number of balls finally placed in urn i, then $M_i \sim K(\alpha_i, \beta)$, and all of the M_i are independent of each other, $i = 1, \ldots, n$. □

Proof: Let $\mathbf{M} \equiv (M_1, \ldots, M_n)$, and $\mathbf{m} \equiv (m_1, \ldots, m_n)$. It will be shown by induction that $P(\mathbf{M} = \mathbf{m} | M = m)$ is given by (24) for $m \geq 0$. Since (24) is equivalent to (23), the theorem follows upon removing the condition on M.

Let $Q(\mathbf{m})$ be $P(\mathbf{M} = \mathbf{m} | M = m_1 + \ldots + m_n)$, and note that $Q(\mathbf{0}) = 1$, a special case of (A18) where $m_1 + \ldots + m_n = 0$. Suppose $Q(\mathbf{m})$ is given by (A18) for all $\mathbf{m}$ such that $m_1 + \ldots + m_n = k$; let $\mathbf{e}_i$ be an n-vector all of whose components are zero except for component i, which is 1; and let $k_i = m_i - 1$, $i = 1, \ldots, n$ (if $k_i < 0$, then the corresponding

term may be omitted from (A20) below). Then, conditioning on the ball configuration after k balls have been placed,

$$Q(\mathbf{m}) = \sum_{i=1}^{n} Q(\mathbf{m} - \mathbf{e}_i) \frac{\alpha + \beta k_i}{\alpha + \beta k} \qquad (A20)$$

where $\mathbf{m}$ is now any configuration such that $\sum_{i=1}^{n} m_i = k + 1$. $Q(\mathbf{m} - \mathbf{e}_i)$ on the right hand side of (A20) is by assumption given by (A18), and it is now only a matter of some algebra to conclude that $Q(\mathbf{m})$ on the left hand side is also given by (A18). Since $\mathbf{m}$ is arbitrary except for its sum, this completes the inductive proof. $\qquad\square$

<u>Comment</u>: When $\beta = 0$, equation (A19) makes $p_i = \alpha_i/\alpha$ for every ball. The fact that a Poisson random variable produces independent Poisson parts when partitioned in this manner is well known (e.g. Ross (1993)). When $\beta \neq 0$, if each ball is placed in urn i with probability α_i/α, instead of according to (A19), then by Theorem 3 $M_i \sim K(\alpha_i', \beta')$, where $\alpha_i' = \alpha_i/(1 - \beta(1 - \alpha_i/\alpha))$ and $\beta_i' = \beta(\alpha_i/\alpha)/(1 - \beta(1 - \alpha_i/\alpha))$. $E(M_i)$ is still $\alpha_i/(1 - \beta)$, but it is not true that $M_i \sim K(\alpha_i, \beta)$, and furthermore $M_1, \ldots, M_n$ are not mutually independent. These latter properties require that the balls be allocated according to (A19).

REFERENCES

Feller, W. 1957. *An Introduction to Probability Theory and its Applications* (2nd ed.), Wiley, p. 48.

Glazebrook, K.D and Boys, R.J. 1995. "A Class of Bayesian Models for Optimal Exploration", *Journal of the Royal Statistical Society (B)*, **57**, No.4, pp. 705-720.

Horrigan, T. J. 1973. "Stopping Distributions and Stopped Distributions: A Basis for Mine Warfare Measures of Effectiveness", *Proceedings of the Sixteenth Technical Conference of the Naval Minefield Community: the Family of Underwater Weapons*, NOLTR 73-73, US Navy.

Janardin, K. and Patil, G. 1970. "On the Multivariate Polya Distribution: A Model of Contagion for Data with Multiple Counts", *Random Counts in Scientific Work II. Random Counts in Physical Science, Geo-Science, and Business* (ed. G.P. Patil). The Pennsylvania State University Press, pp. 143-162.

Jelinski, Z. and Moranda, P. 1972. "Software reliability research", in *Statistical Computer Performance* Evaluation, (W. Freiburger, ed), Academic Press, pp. 465-484.

Johnson, N. and Kotz, S. 1969. *Discrete Distributions*, Wiley, NY, p. 36-43.

Johnson, N. and Kotz, S. 1977. *Urn Models and Their Applications*, Wiley, NY, p. 194.

Katz, L. 1965. "Unified Treatment of a Broad Class of Discrete Probability Distributions", in *Classical and Contagious Discrete Distributions* (G. Patil, ed.), Statistical Publishing Society, Pergamon Press, Oxford.

McCurdy, M. 1987. "A Cognitive Planning Aid for Naval Minesweeping Operations", USCINCPAC Technical Report (1988 revision), Camp Smith, HI.

Odle, J. 1977. "Minefield Analysis for Channelized Traffic", NSWC/WOL TR 77-109, Naval Surface Weapons Center, White Oak, MD.

Paes, J. 2001. "Javamix: A Tactical Decision Aid to Evaluate Minefield Clearance Plans", Naval Postgraduate School MS thesis in Operations Research, June 2001.

Ross, S. 1993. *Probability Models* (5th ed.) Academic Press, NY, pp. 216-217.

Stone, L. and Corwin, T. 1995. "NODESTAR: A Nonlinear, Discrete, Multiple-target, Correlator-tracker: Part 1", *U.S. Navy Journal of Underwater Acoustics*, **45**, No. 3, pp. 1-16.

Wagner, D.; Mylander, C. and Sanders, T. 1999. *Naval Operations Analysis* (3rd ed), chapter 10, Naval Institute Press.

Washburn, A. 1995. "MIXER: A TDA for Mixed Minefield Clearance", NPSOR-95-011PR, Naval Postgraduate School, Monterey, CA 93943.

Washburn, A.R. 1996. "Katz Distributions, with Applications to Minefield Clearance", NPS-OR-96-003, Naval Postgraduate School, Monterey, CA, 93943.

Washburn, A. 2005. The tutorial "Notes on Mine Warfare", the workbook *Minewar.xls* and a zipped version of the MIXER tactical decision aid prototype can all be downloaded from the "downloads" section of http://www.nps.navy.mil/orfacpag/resumePages/washbu.htm ,as of February, 2006.